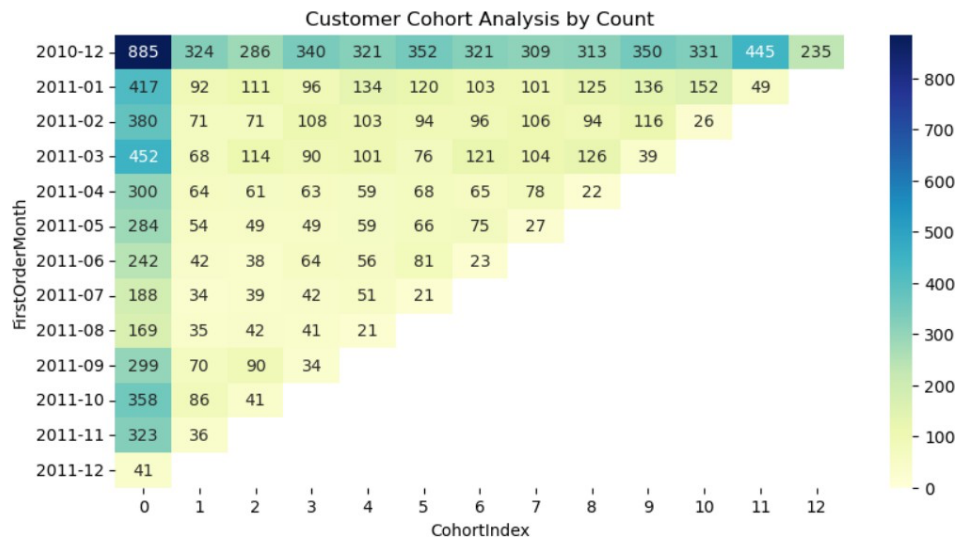


# Cohort Analysis Report

임은석

## 1. Visualization 1

### a. 시각화 그래프



### b. Python code

```
cohort_counts = (
    df1.groupby(['FirstOrderMonth', 'CohortIndex'])['CustomerID']
        .nunique() # 각 그룹의 고유 고객 수(CustomerID)를 직접 계산
        .unstack() # 그룹화된 데이터를 피벗 테이블 형식으로 변환
)

plt.figure(figsize=(10, 5))
plt.title('Customer Cohort Analysis by Count')
sns.heatmap(data=cohort_counts, annot=True, vmin=0, fmt=".0f", cmap="YlGnBu") # annot: 셀에 값 표시
plt.show()
```

### c. 시각화 그래프 설명

- 각 행은 특정 첫 주문 월(FirstOrderMonth)을 나타내며, 각 열은 첫 구매 이후 다시 구매할 때까지 걸린 개월 수(CohortIndex)를 나타낸다. 따라서, 히트맵 안의 값들은 첫 주문 후 다시 구매한 고객의 수를 의미한다.  
예) 2010 년 12 월에 첫 주문을 하고, 2010 년 3 월에 다시 주문을 한 고객은 340 명이다.

#### d. 전략 및 의사결정

- 각 코호트의 첫 달(CohortIndex=0)에서 1 개월 차(CohortIndex=1)로 넘어가면서 고객 수가 크게 감소하고 있다. 첫 1~2 개월 동안 제공하는 특별 할인, 특별 쿠폰, 개인화된 상품 추천 등으로 고객의 관심을 유도하는 전략을 취하면 신규 고객의 초기 이탈을 방지할 수 있을 것이다.
- 2011 년 7 월, 2011 년 8 월을 보면 해당 월의 신규 고객(CohortIndex=0)이 다른 달에 비해 현저히 낮다. 해당 기간에 왜 신규 고객의 유입이 줄었는지 확인해볼 필요가 있으며, 계절별 특화된 마케팅, 신규 고객 혜택 광고 등으로 집중적인 마케팅 캠페인을 실시하여 신규 고객 유입을 지속적으로 이끌어내야 한다.

## 2. Visualization 2

### a. 시각화 그래프



### b. Python code

```
cohort_size = cohort_counts.iloc[:,0]
retention = cohort_counts.divide(cohort_size,axis=0)

plt.figure(figsize=(12, 5))
plt.title('Retention rates')
sns.heatmap(data=retention, annot = True, vmin = 0.0, vmax = 0.5, fmt = '.1%', cmap="YlGnBu")
plt.show()
```

### c. 시각화 그래프 설명

- 코호트 분석을 통해 각 월별 신규 고객 그룹(FirstOrderMonth)이 시간(CohortIndex) 경과에 따라 얼마나 유지되었는지를 백분율(%)로 나타낸 히트맵이다.

### d. 전략 및 의사결정

- 첫 달 이후 급격한 리텐션 감소가 관찰되고, 첫 주문이 2010 년 12 월인 고객들을 제외하면 거의 대부분의 코호트에서 고객 유지율이 30% 이하인 것을 확인할 수 있다. 기존 고객은 시간이 지남에 따라 브랜드와의 관계가 강화된다. 브랜드와의 관계가 오래될수록 고객들의 평균 주문 금액과 구매 빈도는 증가하며, 기업은 고객과의 관계가 강화될수록 안정적인 수익을 얻을 가능성이 높아진다. 따라서, 신규 고객을 확보하기 위한 마케팅 전략도 중요하지만 기존 고객을 유지하는 것이 굉장히 중요하며 신규 고객 유치보다 비용도 훨씬 저렴하다.

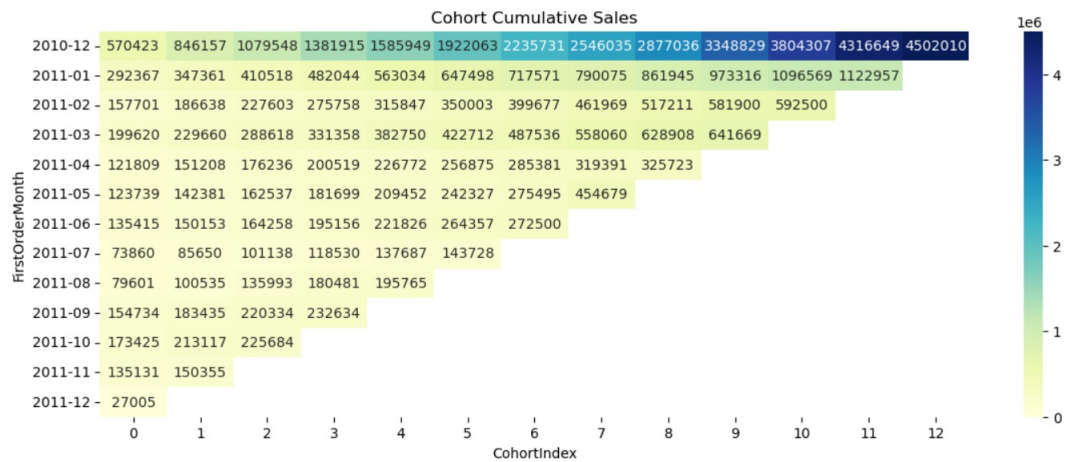
1) 첫 달(CohortIndex=1)의 리텐션을 감소를 막기 위해 초기 사용자 경험(온보딩 과정)을 개선해야 한다. 고객의 구매 패턴에 맞춘 할인이나 맞춤형 추천 제공으로 개인화된 프로모션을 하거나, 제품/서비스 활용법에 대한 튜토리얼 제공으로 고객의 UX 를 높이는 전략을 취한다.

2) 리텐션율 감소가 특정 제품이나 서비스 품질 문제로 인해 발생했는지 점검할 필요가 있다. 탈퇴 고객 대상으로 설문조사나 인터뷰를 실시하여 피드백을 수집하거나 리뷰 데이터와 고객 서비스 문의 데이터를 분석하여 불만 사항을 개선해야 한다. 또한, 리텐션 KPI 를 설정하여

코호트별 리텐션 목표를 설정하고, 각 달별로 추적한다면 초기 고객의 이탈을 줄이고, 장기적인 고객 생애가치를 증가시켜 전체적인 리텐션을 상승을 기대할 수 있을 것이다.

### 3. Visualization 3

#### a. 시각화 그래프



#### b. Python code

```
cohort_cumulative_revenue = (
    df1.groupby(['FirstOrderMonth', 'CohortIndex'])['Sales'].sum()
    .unstack()
    .cumsum(axis=1) # Calculates the cumulative sum across columns (CohortIndex)
)

plt.figure(figsize=(14, 5))
plt.title('Cohort Cumulative Sales')
sns.heatmap(data=cohort_cumulative_revenue, annot = True, vmin = 0.0, fmt=".0f", cmap="YlGnBu") # "BuPu"
plt.show()
```

#### c. 시각화 그래프 설명

- 코호트별 누적 매출(Cumulative Sales) 데이터를 계산하고, 이를 히트맵으로 시각화한 것이다. 값이 클수록 해당 코호트가 지속적으로 매출을 창출했음을 의미하며, 셀 색상은 매출 값의 크기를 시각적으로 구분하여 밝은 색은 낮은 매출, 어두운 색은 높은 매출을 나타낸다.

#### d. 전략 및 의사결정

- 대부분의 코호트에서 초기 매출 기여도가 가장 높고, 시간이 경과하면서 매출 증가 속도가 느려지며, 특정 코호트에서는 3~4 개월 이후 매출이 거의 정체된 것을 확인할 수 있다.

1) 첫 구매 시 할인 쿠폰 제공, 초기 고객 대상 무료 배송 및 번들 할인 프로모션을 하면 고객이 초기 구매를 더 많이 하게 되어 초기 매출 극대화를 노려볼 수 있을 것이다.

2) 시간이 지나면서 매출 증가 속도가 둔화되는 경향이 있으므로, 기존 고객의 재구매를 유도해야 한다. 일정 기간 후 재구매를 촉진하는 프로모션(예: 구매 후 1 개월 내 재구매 시 10% 할인)이나 맞춤형 제품 추천 및 이메일 캠페인을 활용하여 고객의 재구매를 유도하면 초기 구매 후 장기적인 매출 기여도를 높일 수 있을 것이다.

3) 2010 년 12 월에 신규 유입한 고객들의 누적 매출이 다른 월의 신규 고객들보다 훨씬 높기 때문에 해당 시기의 마케팅 캠페인이나 프로모션을 분석하여 성공 요인을 파악할 필요가 있다. 그리하여 성공한 마케팅 전략(광고 채널, 할인 정책 등)을 유사한 시기에 반복 적용하면 매출 증대를 기대할 수 있을 것이다.